CENTRE DE RECHERCHE
SUR LES TRANSPORTS
(C.R.T.)
CENTRE FOR RESEARCH
ON TRANSPORTATION

UNIVERSITÉ DE MONTRÉAL
ÉCOLE DES HAUTES ÉTUDES COMMERCIALES
ÉCOLE POLYTECHNIQUE

LABORATOIRE SUR LA SÉCURITÉ DES TRANSPORTS

LABORATORY ON TRANSPORTATION SAFETY

CENTRE DE RECHERCHE SUR LES TRANSPORTS
UNIVERSITÉ DE MONTRÉAL
C.P. 6128, SUCCURSALE A.
MONTRÉAL
CANADA H3C 3J7

TÉLÉPHONE: (514) 343-7575
EMAIL: crt@crt.umontreal.ca
TÉLÉCOPIEUR (FAX): (514) 343-7121

CRT-752

Study of Operator Performance Measurement : Phase II*

by
Brunilde Sanso[1]
and
Claire Laberge-Nadeau[2]

* This work has been done with Dr. Suzanne Lamarre of SANTRA/LAVALIN

1 École Polytechnique and Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale A, Montréal, Québec H3C 3J7.

2 Département de médecine sociale et préventive et Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale A, Montréal, Québec H3C 3J7.

Centre de recherche sur les transports — Publication #752

January 1991

Dépôt légal – 1er trimestre 1991

91-02144/

ABSTRACT

Human factors have been recognized as important elements of transportation safety. In this paper, we present a literature review on the subject of operator performance measurements for different modes of transportation. The objective of the review was to find tests/devices that could detect or predict impairement while the operator is performing its task.

Keywords: Operator performance measurement, tests/devices, transportation modes, air transportation, air traffic controllers, pilots, metro, train, maritime transportation systems, motor vehicle transportation.

RÉSUMÉ

Les facteurs humains sont bien reconnus comme étant des éléments importants dans la sécurité des transports. Nous présentons ici une revue de la littérature sur des mesures de performance des opérateurs de différents modes de transport. L'objectif de la revue était celui de trouver s'il existe des tests ou dispositifs qui pourraient détecter ou prédire la défaillance de l'opérateur pendant qu'il réalise sa tâche.

Mots-clés: Mesures de performance d'opérateurs, tests/dispositifs, modes de transport, transport aérien, contrôleurs aériens, pilotes, métro, train, transport maritime, transport terrestre.

TABLE OF CONTENTS

Page

1. INTRODUCTION. 1

2. OBJECTIVE 1

3. METHODOLOGY: How the Review was Conducted 2

4. RESULTS 3
   - 4.1 - Air Transportation. 3
     - 4.1.1 - Air Traffic Controllers 4
       - 4.1.1.1 - Psychological and Aptitude Tests and Factors that Could Predict ATC's Performance 5
       - 4.1.1.2 - Physiological Tests 10
       - 4.1.1.3 - Interaction Tests 12
     - 4.1.2 - Pilots. 15
       - 4.1.2.1 - Biomedical Studies. 15
       - 4.1.2.2 - Psychological Studies 17
       - 4.1.2.3 - Interaction Studies 18
   - 4.2 - Metro, Train and Maritime Transportation Systems. 21
   - 4.3 - Motor Vehicle Transportation. 23
     - 4.3.1 - Psychophysiological Studies 24
     - 4.3.2 - New Technological Developments. 29
       - 4.3.2.1 - Tachometers 30
       - 4.3.2.2 - Simulators. 30
       - 4.3.2.3 - The "Yellow Box". 31
       - 4.3.2.4 - Relationship between Simulator Research and Yellow Box Design. 32
       - 4.3.2.5 - Other Technical Developments. 33

5. DISCUSSION. 34

6. CONCLUSION. 36

7. REFERENCES. 40

8. LIST OF ORGANIZATIONS REACHED 49

9. LIST OF INDIVIDUALS CONTACTED 49

1. INTRODUCTION

For a long time, human factors have been recognized as one of the most important elements of transportation safety. In this study, we are concerned with the performance of operators of any type of transportation mode. We are particularly interested in the detection of performance impairment at work as well as in the short-term prediction of that impairment.

The question that was stressed by the transportation authorities in this study of Operator Performance Measurement was whether it is possible to detect a failure in performance due to transitory causes while the operator is working. Another concern was the reliability and the feasibility of these measures. "Can these measures really detect a performance problem, and are the tests involved in the measurement simple, inexpensive and convenient?" Finally, what was the desirability of those tests; "are they intrusive, cumbersome or unethical?"

This work is presented in the following way. The objective and the methodology of the work are briefly exposed. Then the section called Results, in which different trends on Operator Performance Measurement are described. The annotated literature review is included in that part. Then we have the discussion concerning the literature review and, finally, the conclusions and recommendations.

2. OBJECTIVE

The main objective of this work is to provide an annotated bibliography on existing tests/devices related to the measurement of operator performance. We are particularly searching for tests/devices that could detect performance impairment during the task, that is what we call on-site performance tests; we are also looking for tests that could predict performance just before the task is performed, that is what we call a short-term performance test.

A second objective is to do a review on the tests, procedures and apparatus related to operator performance which are in use or are described in the literature. This review will also include long-term predicting tests.

The most pertinent tests will be extracted and commented upon based on their reliability, feasibility and desirability factors in order to identify those which could be practically implemented for our main objective.

3. METHODOLOGY: HOW THE REVIEW WAS CONDUCTED

The review is presented under different perspectives. Although many studies were obtained in the "classical" manner (data base interrogation, articles, books), we have also relied on personal contacts to identify the tests that are now in use in different modes of transportation.

We collected the information in the following ways:

3.1 - Publications were reviewed in English, French, Spanish, Italian and German;

3.2 - National and International Agencies related to Transportation were contacted;

3.3 - Research Groups or Individuals were identified and contacted;

3.4 - Site visits were carried out.

It is important to point out that this does not represent an exhaustive review of all the material related to operator performance. We would also like to stress that the studies cited in this work are only a subset of all the studies found by the authors, and that after an initial analysis, only the most relevant have been kept: eighty-four papers were retained. Most of them focussed on long-term prediction of performance. Therefore, it was not considered pertinent to carry out a meta-analysis of the studies presented in these papers.

4. RESULTS

We have found no tests that could detect performance impairment during the task. The only apparatus found is a simple device called "dead-man button" that is commonly used in subway systems to detect impairment of the train operator. Although we have found a great amount of interest in such tests in the scientific and commercial community, we were surprised to realize that there were no implemented tests or devices to detect performance impairment during the task in the literature. Therefore no annotated bibliography on on-site performance tests is possible. Thus, we pass to our second objective: to conduct a literature review on the subject.

The presentation of the literature review is divided in three subsections: the first one corresponds to Air Transportation, the second to Subway, Train and Maritime Transportation, and the last to Motor Vehicle Transportation (cars, buses and trucks).

4.1 Air Transportation

Two main groups related to Air Transportation safety are the Air Traffic Controllers (ACTs) and the Pilots. They are also among the most studied operators in transportation.

The first agency contacted to review existing tests for pilots and ATCs was the International Civil Aviation Administration ICAA (see list of individuals and organizations) and, in particular, Dr. S. Finkelstein, Director of Aeronautical Medicine. At the same time, a CBC television program on Air Traffic Safety ("A sky full of planes") presented in David Suzuki's "The Nature of Things" was reviewed and its producer, Italo Costa, contacted.

The personal contacts as well as the reviewed publications pointed to the studies carried out by the FAA Civil Aeromedical Institute in Oklahoma. In their facility in Oklahoma a multidisciplinary group of researchers have carried out different studies over the years to evaluate or

predict pilots and ATC performance. A meticulous review of the FAA's and other publications in Air Transportation safety have brought us to identify three major aspects of research for operator performance: 1) psychological or aptitude factors for long-term performance; 2) biomedical or physiological factors involved in operator performance; 3) man-machine interaction tests which could detect on-site performance impairment.

It is interesting to point out that there is a "chronological concern" in the studies carried out over the years on ATC performance. It is clear that researchers were first concerned with psychological factors, then with physiological factors and just recently with the interaction tests. This "chronological concern" is less evident in the review of pilot performance studies. In fact, for pilots, one could say that earlier studies were mostly concerned with biomedical factors.

For the ATC and the pilot literature, the review will be presented by the type of factor under study, each type being a sub section. Within each subsection, the studies are presented by date of publication.

4.1.1 Air Traffic Controllers

In the study of ATC performance, the earlier studies on psychological factors affecting ATC performance were mainly concerned with prediction: who is the ATC individual and which of the ATC trainees will be good ATC professionals; in other words, what personality factors affect ATC performance?

The second group of studies corresponds to the physiological factors that could predict or evaluate ATC performance. It is interesting to point out that we go from the "long-term performance problem" to the notion of "short or on-the-spot performance". The major concern is not only to predict which trainees will be successful professionals, but also how a stressful environment can act on the individual and how this effect can physiologically be measured.

Finally, the interaction tests evoked the idea of predicting performance on a long-term basis. However, the most recent studies have been carried out with regard to an imminent change in the ATC task. We will discuss this change in the subsection where we review the interaction tests.

4.1.1.1 Psychological or Aptitude Tests and Factors that Could Predict ATC's Performance

The concern for tests or indicators that could predict ATC performance go back to the late 1950s and early 1960s. In [1] through [6] it is shown that aptitude tests, experience related to air traffic control, and age of entry into training are important factors in the future performance of Air Traffic Controllers.

One of the first studies carried out on suitable procedures for the selection of ATC was a joint research project planned by the Center for Civil Aeronautics Administration (now Federal Aviation Agency Aeronautical Center) in 1957. The results of this project, reported by Brokaw [7], indicated that psychological tests are good predictors of the academic grades and instructor ratings, which are obtained by the trainees just after their training.

In 1961, job performance evaluations and medical history data were obtained from 149 out of the 197 men that were part of the study reported by Brokaw. Trites [8], who reported this follow-up study, considered some predictor variables that were taken into account in the future project. The conclusion of his study can be summarized in four points:

(1) psychological tests can make a useful contribution to screening applicants for air traffic control work;

(2) instructors, in the ATC school, can make exceptional predictions on the future job performance of their trainees;

(3) older trainees tend to receive poorer job performance ratings as the years go by;

(4) future medical history cannot be predicted by early psychological tests.

Gradually, we observe in the subsequent studies on performance evaluations, that the ATC work involved subspecialties.

To understand the other studies, we believe that it is worthwhile to digress briefly to explain these subspecialties: Terminal (Tower) Traffic Control, Air Route (Enroute) Control and Approach Control.

Tower ATC are controllers who give permission to pilots to land or to take off. A few miles after the take off, the Enroute or Air Route controllers take the control of the plane. Those controllers monitor the airplane until it gets near the airport where the Approach controllers and then the Tower controllers take over.

Enroute controllers can be Flight Information Region Controllers (or Flight Service Station Controllers) or Controlled Area Center Controllers. Flight information Region Controllers monitor the plane without having the plane on radar and air traffic control in their region relied only on information given by the pilot. In those regions, pilots are in full control of their plane; controllers can only advise, and do not have authority over the pilots. On the other hand, in those air corridors where traffic is very heavy (on the American East Coast for instance), the airplane is on radar-control during its entire flight. Those regions are called Controlled Areas and ATCs for those areas have full control and authority over the plane.

When the plane is about 30 miles away from the destination airport, the Approach controllers take over from the Enroute controllers. They monitor the airplane until it is ready for landing; then the control is passed on to the Tower controller.

In 1965, another paper of a series of seven papers identifying problems in air traffic management was published [9]. The paper reported the results of a study whose aim was to identify and evaluate the potential of aptitude tests for selection of Tower ATCs. This study indicated that ATC performance could be well predicted by a composite of four tests measuring the following:

(1) numerical ability;

(2) non-verbal abstract reasoning;

(3) ability to solve simplified air traffic problems;

(4) verbal abstract reasoning.

Experience related to air traffic control was found to be of importance, but most of all, the verbal abstract reasoning aptitude was found unique for the prediction of Tower ATC's performance. It is interesting to point out that previous research had failed to identify the importance of verbal abstract reasoning in the performance of Enroute controllers. In general, the study concluded that, with the exception of verbal abstract reasoning, aptitude tests available at that time provided adequate measures to predict future performance. It was then recommended to include verbal abstract reasoning tests in the battery of tests necessary for ATC selection.

A seventh and final report on problems in air traffic management [10] was concerned with the measurement, structure and prediction of job and training performance of ATC specialists. One of the conclusions reached by the study is that job supervisors can make subjective evaluations of ATC's performance that are reliable and predictable; however, evaluation of Tower ATC's performance are more reliable and predictable than evaluations of Enroute ATC's.

The study did not determine whether the lower level of prediction among Enroute ATC's could be accounted for the low reliability of the Enroute supervisory ratings or for the nature of the Enroute task itself.

In 1968, results from a study conducted by Cobb [11] were reported. The study focussed on the experimental use of commercially-published aptitude tests and the determination of their validity as predictors of performance. The tests were taken by over 300 Marines and 600 Naval students who were enrolled in an ATC training course. The tests chosen were selected from a group whose validity had been established in previous research carried out on 400 FAA trainees. The results indicated that only four of the numerous tests could be retained. Regarding the differences in performance between the two groups in the study, the Marines were found to have a lower failure rate that the Naval trainees. Since the Marines were also found to have been selected from a group with higher military-screening-and-classification (MSC) test scores than the Naval trainees, the study suggested that ATC trainees should be selected from those who presented higher scores on the MSC tests.

In the late 60's and early 70's, the FAA was particularly concerned with the relationship between personality factors and performance. Trites, Kurek and Cobb [12] studied the effects of over and under-achieving during the training program on job performance, and also studied the relationship of performance to the scales of the California Personality Inventory [13] and found significant relationships. However, in 1969, Karson [14] was unable to find a relationship between supervisor ratings and personality variables. Since this finding was surprising, in 1970, Karson and O'Dell [15] attempted once again to find the relationship between some personality variables and job performance. However, their study did not show significant results.

In 1971, a new study by Cobbs [16] was published, in which, again, the main concern was the determination of relationships between aptitude scores in seven commercially published aptitude tests and training-course performance. Results indicated that a composite score involving only four of the tests could have been used with considerable effectiveness to predict military ATC performance. In 1972, Cobb and Matthews published another study [17] on the development of a new aptitude test called "Directional Heading" (or DHT) for the selection of Air Traffic Controllers. This test required individuals to rapidly interpret letters, symbols and degrees in order to determine directional headings. The test was considered a successful instrument for predicting which ATC trainees will pass their training courses.

To end the review on the psychological factors and ATC performance we will comment on an interesting study by Smith [18]. In this work, an overview of research findings on ATC is presented. Results regarding personality, motivations, aptitude and attitude with the notion of "success" in the air traffic profession are shown. Data from former studies were assembled and compared with data from a general male population and from a sample of pilots. The scores of ATCs differ very little from those of the general population, and only three out of sixteen scores fall outside the range that could reasonably be expected in the general population. Those scales are: 1) intelligence; 2) a scale reflecting a tough-minded and self-reliant approach of life; and 3) a measure of self-control and decisiveness. Interestingly enough, the controllers were, in general, found closer to average males than were the pilots. In other tests, it was found that controllers tend to do well on measures of spatial orientation, spatial visualization, memory, numerical ability and abstract reasoning. In

the same study, a survey carried out on 792 ATCs was able to assess that what controllers like the most about their profession is the work itself, and what they hate the most are administration and working conditions. The overall satisfaction of ATC's in being in the air traffic control profession was 91% whereas a typical percentage of satisfaction of an occupational group is 80%. In general, the study concluded that ATCs are bright, action oriented, intolerant of routine, and demanding of opportunities to actively participate in the decision-making process.

4.1.1.2 Physiological Tests

Between 1960 and 1963, a large number of ATC students in training at the FAA Aeronautical Center in Oklahoma City underwent biomedical evaluation conducted by the Civil Aeromedical Research Institute (CARI). The study, whose findings were published in 1971 [19], concluded that:

1) physical fitness level in ATC was comparable to general population norms and;
2) expected physical fitness parameters decreased with age. The age-related decreases were generally parallel with the ATC training failure rate.

In the seventies, a group of scientists at the FAA tried to measure and evaluate what we usually call "stress". Stress is a factor greatly identified with the ATC profession. The popular belief is that all ATCs are usually under a high level of stress. In 1974, Melton et al. published a report [20] in which they presented a new index of stress that integrated various well-known indicators of stress. This index, called "C", is useful

only within a given population that is exposed to the same stressors (e.g. the ATC population). The index is calculated from the urinary content of 17-ketogenic steroids (st), epineprine (e), and nor epinephrine (ne). In the first report by Melton et al., it was found that the principal stressor at the Opa Locka Tower control center (OPF) was the heavy volume of air traffic. Values of C_s when ATCs were off duty were low.

Another study by Melton et al. [21] compared the stress index on two different groups of ATC. The first group was at Atlanta Air Route Traffic Control Center (ATL) and most of the ATCs worked on a 5-day shift rotation. The second group worked at Forth Worth Air Route Traffic Control Center (FTW) mostly on a 2-2-1 schedule. On the 5-day schedule, the controller works 5 days on the same shift, has 2 days off, then works 5 days on a different shift. On the standard 2-2-1 rotation pattern, the controller works a different shift every day. A controller on the 2-2-1 schedule has 80 hours off-duty between workweeks as compared to 72 hours on the 5-day schedule. However, the extended off-duty period comes from diminishing the off-duty period during the week. The study concluded that physiological stress at (ATL) was significantly greater than at (FTW). The difference could probably be accounted for the greater air traffic volume at the (ATL) facility and not by the shift schedules.

Melton et al. [22] published a study in 1978 in which the same thirty-two subjects who had participated in previous studies from 1969 to 1973 were re-studied. Their stress index was measured and compared with former data. It was concluded that the active controller group showed an evidence of reduced chronic stress. Possible interpretations were greater

experience in their jobs, less stress at their new facilities, improvements in the entire traffic control system, and the normal effects of aging.

The last "physiological" study to be presented includes physiological, psychological and interaction tests. In this study [23], forty-five male subjects performed a task for one hour in a simulator. The task consisted in responding, as quickly as possible, to infrequent changes in alphanumeric symbols. At the same time, some physiological recordings were made: blood pressure, oral temperature, skin conductance, body movements, heart rate and heart-rate variability. In addition, performance rates such as response time were obtained. The subject also recorded his or her personal feelings of boredom, irritability, attentiveness and fatigue. Two groups were formed according to their boredom-monotony rate. One interesting finding of this study was that the high boredom-monotony group showed greater increase in "long response time", heart-rate variability, strain and a greater decrease of attentiveness.

4.1.1.3 Interaction Tests

In this section, we review the studies that were concerned with the interaction between the Air Traffic Controller and the machine in which he or she performs his or her task.

The main objective of nearly all of the studies reviewed was to predict, by means of a simulator, which ATC or ATCs trainees will be most efficient. Since the late 1970's, there has been a lot of discussions about changing the ATC task from an active to a passive task of observing a highly automated computer controlled system. This fact has increased the interest in testing the ATC while interacting, in different ways, with simulators.

In 1965, the first report by Pearson et al. [24] describing the development of a radar air traffic control research task appeared. In this study, a system is described in which various elements of the radar ATC's task are presented repeatedly. The system was programmed in such a way as to test operator skills such as vigilance, problem solving, monitoring and reaction time.

Later, in 1972, a new FAA publication [25] showed the results of five different studies in which 229 ATC trainees were tested on a Multiple Task Performance Battery (MTPB). It was concluded that the MTPB approach to selection offered promise as a screening device for ATC applicants but that it was necessary to do further research.

A different type of study was published in 1978 [26]. As the role of the controller is gradually changing from that of an active participant to an overseer of a complex computerized system, the FAA has been mainly concerned with the ability of an individual to sustain a high level of attention to a passive task. It was thought that the gender factor of the ATC could be related to the ability to sustain attention during monitoring performance. The study, which tested 26 ATC men and 26 ATC women over a 2-hour session, concluded that detection latency to the signals increased significantly during the session, regardless of the sex of the controller. Apparently, the alertness decline was equal for both sexes.

In the late 70's, the FAA constructed a Radar Training Facility (RTF) in Oklahoma city to aid in the screening of appropriate personnel in air traffic control. In [27], results are shown on the differences in screening between over-the-shoulder and computer-derived scorings. The study concluded that computer-derived measures are far more reliable in predicting ATC performance than over-the-shoulder measures.

Another study [28] examined the possibility of measuring predictors of complex monitoring performance. The idea here was to try to predict which ATC would do well in a more passive highly automated environment. The study was not conclusive as to effective predictors of performance.

In another study by Thackray and Touchstone [29] the question of boredom and alertness of an ATC facing a more passive role was again addressed. The study also included non-ATC's alert response to a radar monitoring task. It was found that subject-controlled conditions were, in general, superior to computer-controlled conditions and that target detection time was significantly shorter in an ATC than in an non-ATC. Both groups, however, showed a similar increase in target detection time after 90 minutes of task performance. The results suggested that 90 minutes may be the maximum time that attention can be sustained in the type of task described.

In 1986, Thackray and Touchstone [30] reported an experiment which examined the possible relationship between the coronary prone type "A" behavior and performance on a complex monitoring task. Thirty-six male subjects, half classified as type "A" and half as type "B", monitored the radar for a 2-hour period. There was no difference between the groups regarding task performance, nor any significant physiological difference.

Finally, in 1988 [31], a new experiment conducted by Thackray and Touchstone [31] was reported. A simulated ATC task was developed to study the effects of prolonged monitoring. Two types of events were employed: the readily detectible loss of altitude information in alphanumeric symbols and the detection of two airplanes at the same altitude on the same path. The results revealed that time to detect the second event increased significantly over the monitoring period.

4.1.2 Pilots

As we have mentioned above, the Pilot performance studies can easily be subdivided into physiological, psychological and interaction studies. However, the "chronological concern" is not as evident in the Pilot review as it was in the ATC review.

4.1.2.1 Biomedical studies

One of the first studies reporting physiological monitoring of pilots was by Melton and Wicks [32]. In this study, records of heart rate, lateral eye movements and vocal interchange between student and instructor were taken throughout a conventional private pilot syllabus. It was found that heart rate was higher before solo than before duo flights, that the frequency of lateral eye movement was higher in solo flights and that the number of words spoken were useless as a predictor of success in the program.

Another physiological study was reported in 1968 by Vaughan, Higgins and Funkhouser [33]. In this study a group of men were exposed to different environmental temperatures and were asked to perform tasks related to pilot manual dexterity, as well as to take a written motor coordination test. It was found that coarse hand movements are independent of body thermal state but that more discrete tasks involving finger dexterity and motor co-ordination can be most efficiently performed in warmer environments.

Regarding the use of tranquilizers in flight trainings, in 1969, a report was published by Melton, Hoffman and Delafield [34]. In the study, male subjects were given flight training. Half of them received chlordiazepoxide (librium) and the other half received a placebo. The treatment

was given only on solo flights. Differences could not be distinguished between the two groups regarding the urinary excretion of cathecholamines and corticosteroids or the heart rate recorded in flight.

In 1972, a study by Billings et al. [35] was reported. In this study very experienced aviators flew in a simulator under the influence of alcohol at different levels: 40, 80 and 120 mg of blood ethyl alcohol. As one could expect, the more experienced pilots maintained their ability to guide the aircraft, even with a higher alcohol level, better than the less experienced pilots. However, both groups demonstrated progressive increases in procedural errors. Therefore, it was concluded that even very low levels of alcohol exert a negative effect on performance particularly for inexperienced pilots.

The effect of high cockpit temperature on physiological responses and performance was studied and reported by Iampietro et al. [36]. In this study, the pilots flew a simulator while exposed to cockpit temperatures varying from 25°C to 60°C. Performance was scored as the number of deviations from a predetermined path. As expected, there was a definite decrement in performance as the temperature increased.

A final report on biomedical factors related to safety was a study of the effects of ethanol on visual activity [37]. Although the study was performed on pigeons, the conclusions reached are in accordance with the literature regarding the effects of ethanol on traffic safety. Those conclusions were: (i) low doses of ethanol may seriously impair peripheral visual functions, (ii) effects are highly dose dependent, (iii) effects of low doses of ethanol may not be extrapolated from high-dose effect, the effects of the same dose of ethanol may vary widely among the individuals.

4.1.2.2 Psychological studies

In 1969, a study by Pearson and Thackray [38] appeared which attempted to predict performance from response to attitute questionnaires. In this study, based upon responses to an item on fear of shock in a questionnaire describing numerous stressful situations, subjects were classified into two groups. After training, individuals in both groups were informed that they would be required to maintain their training performance level or be shocked (stressing situation created). The results indicated that those individuals that did not fear shock performed much better under the stressing situation created than did the individuals of the other group. Therefore, it was concluded that attitude questionnaires may be used to predict performance.

Yanowitch et al [39] published, in 1972, a new approach to investigate aviation accidents. This approach, called "psychological autopsy" enabled a retrospective portrait of the pilot-in-command following to an accident. The approach gave investigators insight into the situational stress that could have diminished pilot performance like a divorce or a recent loss in the family. Another paper related to "psychological autopsy" appeared in 1973 [40]. In it, it is explained that the relationship between the pilot and his aircraft is such that the aircraft may be thought of as an extension of the pilot himself during the act of flight. The pilot is viewed as a competent individual whose work is related to stressful situations. However, the study concluded, if an additional stress is accumulated, the aircraft may be seen as an instrument of self-destruction.

One concern in the Aviation Community has been to determine the psychophysiological effects of aging. In 1978 [41], a study determining the factors associated with pilot performance and proficiency were identified. The study was based on three methodological approaches considered representative of pilot performance:

(a) the analysis of successful pilot behavior under simulated and real conditions;
(b) the analysis of pilot errors related to aircraft accidents;
(c) the evaluation of pilot performance during selection and training.

The study identified 14 factors related to pilot performance: (1) perception, (2) attention, (3) reaction, (4) orientation, (5) sensorimotor, (6) stamina, (7) cognition, (8) interpersonal relations, (9) decision making, (10) experience, (11) hearing, (12) personality, (13) mechanical ability, (14) motivation. In the study, no attempt was made to assign weights to the mentioned factors.

4.1.2.3 Interaction studies

We include in this section tests related to man-machine interaction. However, it is important to point out that many of these tests integrate physiological and psychological factors as well, especially those related to the effect of alcohol on performance.

The first study we report in this section is the one by Jennings, Chiles and West (1972) [42] concerning a two-dimensional tracking task. In this study nineteen subjects were tested in a complex performance device designed to measure monitoring, information processing, pattern discrimination, and group problem solving. Individuals had to perform a tracking task alone and in combination with other tasks. The tracking task was shown to be reliable as a good indicator of performance in both cases. Also, the

task was shown to be sensitive to work load effects from other tasks and also to have an effect on some of the other tasks.

In 1973, a study by Lewis [43] was published. In it, the frequency of anti-collision responses were evaluated as a function of traffic density, ATC warnings and competing behavior. Eighteen pilots had to fly in a simulator for two hours. During that time, the traffic, the ATC warnings and the ATC clearances were varied. The pilot had to face a target which remained constant during the flight. In order to observe the target, the pilot had to press a button; if there was no traffic, there was no stimulus; if traffic was present, a burst of flashes became visible. Two values of traffic frequency were programmed as well as two values of ATC warning frequency. The results showed that performance was affected by the increase of ATC clearances and from traffic warnings.

Also in a 1973 study [44], pilots performance was evaluated while using two different types of aircraft position indicators. The traditional indicators showed a moving horizon and the plane in a fixed position. The experimental indicator showed a moving plane against a fixed horizon. The study was conducted on two groups of civilian pilots classified as "experienced" and "inexperienced". It was shown that, for experienced pilots, performance was unaffected by the new indicator design. However, the unexperienced group showed better performance with the experimental indicator design.

Another study related to different indicator systems was published in 1979 [45]. In it, simulator experiments were conducted to compare pilot performance when faced to four different types of glidepath indicator systems in the severely visually reduced nighttime environment. It was found that differences in performance with different indicators were

attributed to the rate of information change provided by the given system and to the rate of observing the indicator during approaches.

We will now review some of the studies whose objective was to assess alcohol effects in relation with the interaction tests. In a 1978 study [46] eleven pilots (7 men and 4 women) were trained on a Multiple Task Performance Battery (MTPB), static and dynamic tracking, slope instruments, and a speech intelligibility test. Subjects were given tests in the evening, then drank prepared beverages (alcohol or placebo), and were tested again. After 4-5 hours of sleep, they were given breakfast and were tested again. Results showed that immediatly following drinking, performance on the MTPB and the tracking task clearly deteriorated. "Hangover" tests presented lower scores than "evening" tests, but there was not a clear indication that performance was greatly diminished. It was also found that speech perception scores were unaffected by alcohol.

Another study relating alcohol, hangover and performance was reported by Collins [47] in 1979. In it, eight pilots were trained to perform on a two-dimensional tracking task (joystick control) and to respond as quickly as possible to the onset of a red light by depressing a button on the joystick. The experiment followed the same pattern described in [46] except that there were ground-level sessions and simulated altitude sessions. Again, impairment of tracking performance and visual reaction time was found during the period just following alcohol ingestion. No hangover related performance impairment was found, nor any significant attitude/ alcohol performance interaction.

Another interesting study on alcohol effects was reported in 1983 [48]. In this experiment fifteen nonalcoholic men drank, in the evening raising their alcohol concentrations to 100-200 mg/dl for at least 2 hours. After seven hours of sleep, they ate breakfast and then completed a battery of tests. Some physiological measures were taken. It was found that the

urinary excretion rates of catecholamines and ketogenic adrenal steroids were increased by alcohol. Tests affected by alcohol were: number comparison, number addition, analysis of complex statements and adaptability. Tests unaffected were abstract reasoning, digit code, digit symbol, short-term memory, hand steadiness, pursuit tracking, road-and-frame.

The last experiment reviewed regarding the effects of alcohol was the study reported in 1988 by Collins and Mertens [49]. Again, trained men performed on the MTPB with and without alcohol. Subjects breathed appropriate gas mixture through oxygen masks at both ground level and altitude. Younger subjects performed better than older subjects but performance of both age groups was significantly affected by alcohol. Again, no significant results were obtained regarding differences in performance at ground level and at higher altitudes.

4.2 Metro, Train and Maritime Transportation Systems

If in the air transportation system there is an abundance of studies related to operator's performance, in subway, train and maritime transportation systems, there was very little information that the authors of this study could find.

For the subway system one could explain the lack of studies by the "close loop" nature of the system. Besides, since subway systems are almost completely automated in some cities, there is very little place for human error. In fact, the safety concern in the railway community is translated more into "safety directives" for signalling than into the evaluation of operator performance.

Very little was found for the Maritime systems, except, again, concern for safety directives and the typical pre-employment or annual medical check-ups.

For a subway system, the S.T.C.U.M. (Montreal public transport system) was contacted and their operator's performance evaluation discussed. The result of the discussion can be summarized as follows:

1. The operator has to pass a complete pre-employment medical exam and repeat it every two years (or yearly depending on the age of the operator and other conditions).
2. No psychological tests are required by the S.T.C.U.M.
3. A periodic knowledge up-to-date test is taken every three years.

Regarding the tests, measures, systems, or apparatus monitoring performance, since the Montreal subway is almost completely automated, much emphasis is put on the alarms and back-up control systems of the train. However, there is a system called "dead-man" which consists in a button that must be pressed by the operator while the train is running. In case of a serious operator impairment, the lack of pressure on the "dead-man" button will send a signal to the Metro Control Center. The Control Center will then try to make contact with the operator and, in the case of no response, help will be sent immediatley.

For Maritime systems, several organizations and individuals were surveyed. First, we contacted the IMO (Intenational Maritime Organization), the "Pilotes de la voie Maritime du fleuve St-Laurent", and the coast guard division of Transport Canada. In addition, one Venezuelan and one English oil ship captains were interviewed by telephone. The results of these contacts can be summarized as follows:

1. International conventions do require that ship operators pass fitness and medical tests at least every five years.
2. No specific psychological tests are required.

3. There is a radar warning device, ARPA, that can help the ship operator in navigation. However, no operator performance detection system (such as a breathalizer) exist inside the ship.
4. The captains interviewed were not opposed to the idea of having a device that could detect performance impairment installed inside the ship.
5. Both captains thought that the installation of a device such as a breathalizer that could prevent the starting of the navire was feasible.

Regarding railway systems, besides some standard data base interrogation, the following organizations were contacted: OTIF (Office Central des Transports Internationaux Ferroviaires), UIC (Union Internationale des Chemins de Fer) and CN. Except for the international conventions and the standard fitness pre-employment testing, no new material regarding operator performance in these systems was found.

4.3 Motor Vehicule Transportation

Regarding operator performance in vehicular transportation systems (cars, trucks and buses), a very diversified set of sources was consulted. We started with the traditional data base interrogation followed by contacts with international organizations, transportation companies and researchers in the field of transportation. The organizations and companies contacted were IRU (International Road Transport Union), CABANO, PROST TRANSPORTS, and HUNGAROCAMION (Trucking Companies). Among the researchers in the area of transportation, we could discuss with the group of psychologists, led by Jacques Bergeron at the Laboratory of Transportation Safety, in the C.R.T. of the Université de Montréal; they are doing at the moment a research project on psychological factors affecting driver

performance and behavior. We also have contacted a group doing research on a simulator at Concordia University; the aim of this simulator is to train handicapped persons for driving and is not for performance detection. We regret that the on-site visit could not be arranged.

Regarding the studies carried out by international organizations such as the IRU, we should mention that they are very concerned about safety and driver performance. However, much of the emphasis is placed on the driver's preparation preceding employment, on safety measures for trucking companies and on safety evaluations of new devices designed to improve the driver's comfort and his/her safety. All these ideas are underlined in several IRU publications (see [50], [51], [52], [53]).

The rest of the review concerning vehicular systems is divided into two major subsections: psychophysiological studies and new devices.

4.3.1 Psychophysiological Studies

The studies we have reviewed, as well as the personal discussions we have had with experts in the field, led us to believe that, although there has been research related to the psychophysiological aspects of driving a vehicle, some questions have not been completely answered: What are the mechanisms that make some drivers better than others? What makes some drivers take higher risks? How will a particular driver react to a stressful situation and how do external factors affect a driver's behaviour? The answers to those questions still remain as topics for research.

First of all, we mention the recent work by Bergeron and Vaugeois [54] in which an extensive literature review was conducted on the phenomenon of "focused attention" while performing a driving task. This phenomenon can be manifested under the effect of stress or after an increase in mental charge. In his study, Vaugeois reviews the literature by comparing different theories that can explain the phenomenon of focused attention.

Joly and Bergeron [55] state that driving a vehicle is a task in which the individual has a large control over the level of difficulty of the task and, consequently, on the risks he or she is willing to take. Several theoretical models have been introduced to explain a driver's risk behavior. Among those models, two theories emerge: Wilde's theory of risk homeostatis [56] and the theory of resources (see [57]). Wilde's theory states that, by evaluating external factors such as signals, environment, etc., the driver can predict some driving events and that this prediction determines the level of risk he or she is able to take. This theory postulates that subjective and physiological measures related to driving a vehicle are independent of the complexity of the task. On the other hand, the resource theory states that the human operator can be modeled as a reservoir with a finite capacity. The more a central task needs resources, the less there will be resources available for a secondary task. This theory postulates that subjective and physiological measures do vary as a function of task difficulty.

Vaugeois [58] proposes a new research in which both theories will be evaluated as predictors in simulated vehicle driver task. Vaugeois points out that the Taylor's results [59] lead us to believe that some physiological parameters do remain constant while performing a driving task. On the other hand, he stresses with Joly and Bergeron's results that it is not possible to conclude that there is a stability on electrodermal activity and cardiac rhythm regardless of environmental conditions.

Another interesting study was carried out by Bergeron, Audet and Vaugeois [60] in which several factors and theories relating to attention and individual differences are reviewed. Three types of models are described to explain individual differences. Without going into detail of the three types of models that could explain individual differences,

we mention some tests related to one of the models. These tests are of interest because the identification of individual differences in attention could lead to the identification of individual differences in operators' performance. The tests of that model (described by Bergeron et al.) refer to the possibility that the mechanisms of attention are related to the "filtering" of non-pertinent stimuli (see [61], [62], [63], [64], cited in [60]):

a) Test of dichotic hearing

This type of test consists in presenting through earphones two words simultaneously. One word is presented to one ear and the other to the other ear. The individual is asked to report all the words presented or to identify some specific words.

b) Test of selective hearing

In this test, the individual is given two simultaneous messages; however, this time he or she has to pay attention only to one message. He or she has to vocally repeat the message or identify some part of it.

c) Central-incident hearing test

In this test, the experimenter presents a series of cards in which stimuli in two categories are presented. The individual has to pay attention to just one of the two categories, and therefore everything learned from the other category is considered secondary learning. The individual is asked to match the items from the principal and secondary categories.

d) The Stroop test

There are several variations of this test, one of which consists in presenting three different sets of stimuli to an individual. The first set consists of a series of names of color written in black ink. The second set consists of colored rectangles. The

final set includes names of color written in non-matching colors (for instance, the word "red" is written in green ink). The task consists of reading the words in black ink (first set), naming the colors of the rectangles (second set), and naming the colors of the ink in which the third set of words are written. The time to perform the task as well as the number of errors are registered.

Among the psychophysiological studies done in different types of operators, we mention the work of Dr. Kenford Nedd (see list of persons contacted). Dr. Nedd performs physiological stress analyses and stress audit in order to detect and treat stress-related problems. However, his approach is based on stress treatment and not on performance detection on-site.

An interesting study to be mentioned in this literature review is one by Labiale et al. [65]. This study analyzed a driver's behavior related to fuel consumption. Some 23 individuals were asked to drive a pre-specified circuit. Inside the car, a data acquisition system was installed and some driving parameters were recorded. Before driving, the drivers had to pass several psychological tests. The interest we pay to this study is the review of personality and perceptive or cognitive tests related to a driver's performance and not the fuel consumption results. We hereby include this review:

a) Personality tests

The idea behind these personality tests was that a driver's personality affects his or her driving behavior.

a.1) Test 16 PF de R.B. Cattel [66]

This is a set of 187 questions about everyday life (with 3 possible answers). The test allows the identification of 16 primary personality factors.

a.2) The "frustration test" by Rosenzweig [67]

The aim of this test is to reveal different types of reactions the individual could have in facing the frustrations of everyday life. The test presents 24 drawings of frustration situations: two subjects talk about the situation, and one of them expresses his feelings. The tested individual has to write the answer of the other subject with respect to the drawings.

a.3) "The myokinetic psychodiagnostic test" by Mira and Lopez [68]

This test consists in drawing, without seeing what one is doing, several forms in different space planes. Measures are carried regarding deviations of the drawings from realistic ones to evaluate different personality factors.

b) Perceptive and Cognitive tests

These tests aim at analyzing the drivers' perceptive and cognitive behavior.

b.1) Test type A/B

In the study reviewed [65], the version used was that of Bortner [69]. The idea of the test is to distinguish pattern A behavior (impatient, agressive, etc.) from pattern B (the opposite). The subject is given two statements and chooses the one that corresponds better with his/her perception of himself/herself.

b.2) "Locus control" test

This test is based on Rotter's [70] theory of social learning behavior. This theory states that individuals presenting "internal locus" consider that it is possible to control the environment whereas the ones presenting "external locus" believe that they have no control over the environment.

b.3) "The Matching Familiar Figure Test" by Kagan [71]

This test measures the "reflexiveness-impulsiveness" personality dimension. The individual has to select, among 8 variations, the only figure that is identical to the one shown to him.

b.4 The Stroop test

This test has already been described in our report of Bergeron's study.

This closes the review on the tests chosen by the Labiale et al. study done at the INRETS (L'Institut Nationale de Recherche en Transport et Sécurité), previously called IRT (Institut de Recherche en Transport).

The only personality test that we found validated for predicting safety in transportation was the Cattel 16-sp factor test. This test was validated by Hilakisi et al. [72] in a study done on 916 conscripts of the Finnish Guard Battalion. The test was administered at the beginning of their military service and questionnaires concerning traffic fines and traffic penalties were answered at the end of their 11-month military service. The authors concluded that the 16-factor personality test is valuable in the selection of safe drivers.

4.3.2 New Technological Developments

In this section, we review the existing and experimental devices that could be useful in evaluating or in predicting a driver's performance. We point out that some of the devices are not specifically used for safety purposes but they have been included in the review because the authors consider that they have a potential use in performance evaluation. In fact for a device like a tachometer, the addition of pertinent parameters could detect performance diminishment.

4.3.2.1 Tachometers

The tachometer is a device that registers some important driving parameters. In the past, those parameters were registered on graphic cards and the measurements had to be interpreted and compiled by hand. The new digital tachometers have a microprocessor integrated in it and, therefore, data can be transmitted to a central computer where a suitable program extracts the pertinent data.

The authors visited the trucking company CABANO and discussed with the directors, engineers and technicians of the company. The following data on each driver in service are registered by the tachometers: travel duration, medium speed in every travel sequence, number of stops and idle time and other parameters.

4.3.2.2 Simulators

The use of a driving simulator allows researchers to better evaluate different driving parameters with the aim of improving traffic safety. The research group at the Université de Montréal, already mentioned, directed by Jacques Bergeron, has been working on the development of a simulator system for automobile driving (see [75] for details). The simulator is a driving station where the driver interacts with a visual presentation of a film on a screen. The individual sits on a simulated passenger seat that presents the elements of a real car driver environment. The set is placed in an electrically isolated room which eases the measuring of some of the driver's physiological parameters (such as the cardiac rythm, muscular tension, electrodermal conductance). A video camera and a sound system were installed to study the subject's reactions. A PDP-11 computer in a nearby room allows to control the presented stimuli and to register the driver's manoeuvres. With the potentiometers at the steering wheel and at the pedals, the subject can control the speed of lateral movements. The subject

controls its "vehicle" when some risky simulated situations appear on the screen.

As it has been mentioned before, the simulator constitutes an interesting tool to evaluate the mechanism of improvement or diminishment of a driver's performance.

Another simulator found in the literature [76], is a commercially available truck driving simulator made by the company Perceptronics of Michigan. The TT150 simulates a typical truck cabine in which all the elements to simulate emergencies or other situations can be activated. The use of this simulator is to train truck drivers when faced with dangerous driving situations.

4.3.2.3 The "yellow box"

The yellow box is a system still at the experimental stage whose aim is the registration of pertinent data related to the task of driving. The name "yellow box" refers to the well-known black-box in aviation. A review on the feasilibity of the device has already been carried out by a multidisciplinary group of professors of the C.R.T. (Center for Research on Transportation) directed by Claire Laberge-Nadeau (epidemiologist), in a team with Michel Gou (mechanical engineer) and Jacques Bergeron (psychologist). It is worth mentionning that a grant has already been obtained to specify the pertinent parameters to register and to start the design of the device.

The yellow box, on one hand, will be composed of a set of sensors able to detect some performance parameters and, on the other hand, by a memory unit to store data related to the measured parameters. An essential development in the project will be the information restitution system that will allow the researchers to analyze the measurements.

Some of the measurements to be registered in the yellow box are the following:

- vehicle speed and idle times;
- lateral acceleration;
- wheel locking, braking;
- steering wheel position;
- driver's voice and physiological parameters;
- travel information;
- time and kilometers;
- if possible: some environmental measurements and the possibility of registering the presence of alcohol (breathalizers).
- other parameters if an impact occurs.

It is interesting to point out that the yellow box will be an inexpensive device. In fact, after some inquiries we have found (see [77]) a small box, manufactured in Canada, which is a combination of a micro-computer, a clock, a multimeter, a calibrator, and a scanner. This box offers 20 analogue entries and a memory of up to 20,000 data entries. Its dimensions are 5"x5.5"x1.75" and its weight is just one pound. The box is already in use by the Chrysler company for its auto tests and for the truck tests by GM. The cost of the box is under $3,000 CAN.

We mention the "yellow box" in this literature review because we believe that, by studying the parameters involved in driver behavior it will be possible to detect driver performance impairment.

4.3.2.4 Relationship between simulator research and yellow box design

It is important to point out that the "yellow box" research project and the research that is being carried out on the simulator of the Université de Montréal are two complementary research ideas. In fact, both projects depend on the "Laboratoire sur la Sécurité des Transports du Centre de recherche sur les transports" of the Université de Montréal and

both are looking for the mechanisms that may diminish a driver's performance. First, some interesting driving and physiological parameters will be identified at the simulator of the Université de Montréal (we point out that many interesting conclusions have already been reached by Bergeron and his team). Then, the same parameters could be measured, registered and compared in a "real driving" situation by means of the "yellow box". For instance, on the simulator it may be found that voice modification may be an important detector of performance impaired by fatigue, hypoglycemia, alcohol or drugs. Research is needed to assess if the voice can be an interesting performance detector. The reverse procedure will also be attempted, that is, interesting findings of the "yellow box" research are going to affect the research carried out on the simulator.

4.3.2.5 Other technical developments

A joint venture research is currently carried out in several European Countries. One of the projects is called EUREKA "PROMETHEUS". This research is, in reality, a set of several hundreds of proposed research projects, in different fields, whose aim will be to improve the vehicle, traffic regulation and traffic safety. One particular branch of research that will be touched by PROMETHEUS will be traffic safety and, in partilar, how the projected electronic advances can help to better understand the man-machine interface and can improve driver performance. Some material regarding the PROMETHEUS project is found in [78] to [83]. It is a long-term program of several projects that will be extended up to the 21st century. Several publications are in German and information is not easily accessible. However, we consider that this project should be looked at more closely.

5. DISCUSSION

This review of literature has been very deceiving - in terms of finding on-site and short-term performance tests - in fact none of these tests could be found. However, we have identified some interesting research approaches that could be investigated further:

The review has shown that there is a very close relationship between the notions of performance, personality factors, focused attention and stress.

It has been shown in several studies that the levels of cathecolamines increase during stress situations. In a particular study [49], it was also shown that the level of cathecolamines increases after ingesting alcohol. On the other hand, Bergeron and Vaugeois [55] mentioned that the phenomenon of focussed attention, which is detrimental to global performance, is present in both cases: after ingesting alcohol or living under an excessive mental charge. Therefore, there seems to be a psychophysiological parallelism between two performance detrimental situations. Now, a physiological test detecting an increased level of cathecolamines would be, at the same time too intrusive and too extemporaneous. In fact, even though the cathecolamines action is much more rapid than the corticoids hormones, which are also involved in a stress-response, it takes some time before the cathecolamines can be detected in the urine. However, other physiological parameters could be considered reliable and not as intrusive. For instance, the electrodermal conductance, which has been proven to be a reliable measure of stress, could easily be measured in a small battery of tests. However, depending on the subject, there is an existing natural fear of being tested that could interfere in this type of measures. In an experimental study we could evaluate which physiological parameters interfer with the focussed attention phenomenon and design a less intrusive device than electrodermal on urine testing.

Two other possible interesting physiological measures that could be tested is the frequency of change of the cardiac rythm and the voice changes. These measures seems to be related to excessive stress. Vaugeois and Bergeron are presently doing research on the usefulness of this type of measure. Again, the relationship between this measure and the ingestion of drugs or alcohol is clear not yet.

Therefore, we point out that it is possible to associate the phenomenon of focussed attention to either the ingestion of drugs/alcohol or to an excessive mental charge produced by a stressful situation. The importance of this finding is based on the fact that if it were possible to identify this phenomenon in a particular operator, one might not need to do "intrusive tests", the impairment will be detected even if the causing factors are not known.

Unfortunately, we do not understand yet the interaction mechanisms of mental charge, focussed attention and performance impairement. Therefore, our first recommendation is that further research must be carried out in this direction. We identify, in particular, two branches of research: the study of the interrelations between mental charge, focused attention and performance by means of a simulator and by the yellow box. Both studies are necessary because the simulator's research may identify the mechanisms on how stress and performance can be related, but a simulator is much too expensive and cumbersome to be widely available among different transportation companies. Besides, the yellow box is practically the only type of device that could detect an "on site" performance impairment for non-automatized systems.

A second point is related to the question of whether it would be necessary to conceive different tests/devices for different transportation modes or if just one battery of tests/devices with slight changes could be useful to detect performance impairement in any mode. We do not have the

answer to that question. However, there is again the notion of focused attention and mental charge that have to be considered. In fact, if it can be proven that they act in the same detrimental manner on the performance of any transportation mode, then a common test will be sufficient. On the other hand, if it is found that the mentioned detrimental phenomenon have different effects depending on the operator's task, then modifications on the battery of test/devices should be introduced. We recommend that further research be carried out also in this direction.

Finally, it has been seen in air transportation studies that a perceptual- motor control task such as a tracking task either by itself or combined with other tasks be useful in detecting performance impairment. This task could easily be implemented with a modified "video-game" and can be classified as reliable (see [42]), non-intrusive and feasible.

6. CONCLUSION

The objective of this work was to identify short-term or on-site operator performance impairment test/devices for different types of transportation modes.

We found no test/devices that responded entirely to our main objective.

However, general avenues of research were identified after doing an extensive literature review on studies related to operator performance.

We have observed that the most studied mode of transportation is Air Transportation.

It is important to mention that, even in Air Transportation, many studies have been done on the prediction of performance and very little on the on-site detection of a performance problem. Most of the tests are used

prior to the hiring of operators and are generally used to eliminate those individuals that present negative caracteristics for the performance of their job. Almost no tests exist to pick up "positive" individuals.

Interestingly enough, Air Transportation is the safest mode of transportation; without going into a discussion on the reasons behind this safety record, one could infer that it must be related to the studies and the testings done on ATCs and pilots to improve their performance and to detect factors to performance impairement. There is, in fact, a big disproportion on the amount of studies carried out for the different transportation modes. This can be seen in Figure 1 where the proportion of fatalities in Canada for different modes of transportation are depicted. The following question comes to mind: Is the air transportation safety record related to a constant concern on safety as witnessed by the numerous studies found?

The discussion about the disproportion between the number of studies carried out in the safest mode of transportation versus the little amount of literature found in other modes makes us believe that a device cannot replace a global concern in a transportation system.

Finally, we recommend that further research is needed before we can come up with a reliable test/device to detect on-site performance impairment.

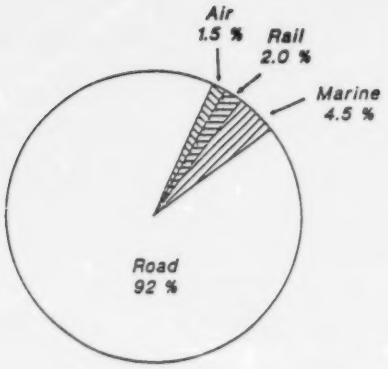

Figure 1 : 1985 Transportation Fatalities in Canada

Source : Road Safety Annual Report, 1987
Transport Canada, Road Safety Divison

This work was done by:

Brunilde Sanso,*Ing., M.Sc.A., Ph.D., and

Claire Laberge-Nadeau,**M.D., M.A.H., M.Sc.,C.S.P.Q.,

in collaboration with Dr. Suzanne Lamarre of SANTRA/LAVALIN.

*Research Associate, Ecole Polytechnique and Centre de recherche sur les transports, Université de Montréal, C.P. 6128, Succursale "A", Montréal, Canada H3C 3J7.

**Professeure titulaire, département de Médecine sociale et préventive, Faculté de médecine. Directrice du Laboratoire sur la sécurité des transports du Centre de recherche sur les transports, Université de Montréal, C.P. 6128, Succursale "A", Montréal, Canada H3C 3J7.

7. REFERENCES

[1] COBB, B.B, (1962), "Problems in Air Traffic Management: II Prediction of Success in Air Traffic Controller School", Aerospace Med. 33:702, 1962. Also published as Civil Aeromedical Research Institute Report 62-2. FAA, Oklahoma City, Oklahoma.

[2] TRITES, D.K., COBB, B.B.,(1962), "Problems in Air Traffic Management: III. Implications of Age for Training and Job Performance of Air Traffic Controllers",Civil Aeromedical Research Insitute Report 62-3. FAA, Oklahoma City, Oklahoma.

[3] TRITES, D.K., COBB, B.B., (1964), "Problems in Air Traffic Management: IV. Comparison of Pre-employment, Job-related Experience with Aptitude Tests as Predictors of Training and Job Performance of Air Traffic Control Specialists", Aerospace Med. 35:428.

[4] TRITES, D.K., (1964), "Age, Aptitude and Experience as Predictors of Performance of Air Traffic Controllers. In : Aviation Psychological Research, Reports of the 5th Conference for Aviation Psychology", Seifert, R. Ed. (Mimeo). Bad Godesberg, Kolner Strasse 70, Germany, pp 157-176.

[5] TRITES, D.K., (1963), "Ground Support Personnel", Aerospace Med. 34:539.

[6] TRITES, D.K., (1964), "Problems in Air Traffic Management: VI. Interaction of Training-Entry Age With Intellectual and Personality Characteristics of Air Traffic Control Specialists", Aerospace Med. 35: 1184.

[7] BROKAW, L.D., (1959), "School and Job Validation of Selection Measures for Air Traffic Control Training." WADC-TN-59-39, Pers. Laboratory, WADD, USAF, Lackland, AFB, Texas.

[8] TRITES, D.K., (1961), "Problems in Air Traffic Management: I - Longitudinal Prediction of Effectiveness of Air Traffic Controllers". Aerospace Med, 32:1112.

[9] COBB, B.B., (July 1965), "Problems in Air Traffic Management: V - Identification and Potential of Aptitude Tests Measures for Selection of Air Traffic Controller Trainees". Report no. AM 65-19, Federal Aviation Agency, Office of Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma City.

[10] TRITES, D.K., MILLER, M.C., COBB, B.B., "Problems in Air Traffic Management: VII-Job and Training Performance of Air Traffic.

[11] COBB, B.B., (Sept. 1968), "A Comparative Study of Air Traffic Trainee Aptitude - Test Measures Involving Navy, Marine Corps and FAA Controllers. Office of Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma City.

[12] TRITES, D.K., KUREK, A., COBB, B.B., (1967), "Personality and Achievement of Air Traffic Controllers. Aerospace Medicine, 38, 1145-1150.

[13] GONGLE, H.G., (1960), "California Personality Inventory Manual", Palo Alto, Col: Consulting Psychologists Press.

[14] KARSON, S., (1969), "Some Relations Between Personality Factors and Job Performances Ratings in Radar Controllers". Aerospace Medicine, 40, 823-826.

[15] KARSON, S., O'DELL, J.W., (Sept. 1970), "Performance Ratings and Personality Factors in Radar Controllers". Office of Aviation Medicine, FAA. Report no. FAA-AM 70-14.

[16] COBB, B.B., (October 1971), "Air Traffic Aptitude Test Measures of Military and FAA Controller Trainees". Office of Aviation Medicine, FAA, Report no. FAA-AM-71-40.

[17] COBB, B.B., MATHEWS, J.J., (May 1972), "A Proposed New Test for Aptitude Screening of Air Traffic Controller Applicants". Office of Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma City, Report no. FAA-AM-71-28.

[18] SMITH, R.C., (December 1974), "A Realistic View of the People in Air Traffic Control". FAA, Office of Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma City, Report no. FAA-AM-74-12.

[19] LATEGOLA, T.M., (March 1971), "Evaluation of Air Traffic Control Students Between 1960-1963". FAA, Office of Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma City, Report no. FAA-AM-71-8.

[20] MELTON, C.E., MCKENZIE, J.M., SALDIVAR, J.T., HOFFMANN, S.M., (December 1974), "Comparison of OPA LOCKA Tower with other ATC Facilities by Means of Biochemical Stress Index, Report no. FAA, AM-74-11.

[21] MELTON, C.E., SMITH, R.C., MCKENZIE, J.M., SALDIVAR, J.T., HOFFMANN, S.M., FOWLER, P.R., "Stress in Air Traffic Controlers: Comparison of Two Air Route Traffic Control Centers on Different Shift Rotation Patterns", Report no. FAA-AM-75-7.

[22] MELTON, C.E., MCKENZIE, J.M., WICKS, S.M., SALDIVAR, J.T., "Stress in Air Traffic Controllers, A Restudy of 32 Controllers 5 to 9 Years Later".

[23] TACKRAY, R.I., BAILEY, J.P., TOUCHSTONE, R.M., (August 1975), "Physiological, Subjective and Performance Correlates of Reported Boredom and Monotory while Performing a Simulated Radar Control Task, Report no. FAA-AM-75-8.

[24] PEARSON, R., HUNTER, C.E., NEAL, G.L., (December 1965), "Development and Evaluation of a Radar Air Traffic Control Research Task", Report no. FAA-AM-65-31.

[25] CHILES, W.D., JENNINGS, A.E., WEST, G., (January 1972), "Multiple Task Performance as a Predictor of the Potential of Air Traffic Controllers Trainees", FAA-AM-72-5.

[26] THACKRAY, R.I., TOUCHSTONE, R.M., BAILEY, J.P., (March 1978), "A Comparison of the Vigilance Performance of Men and Women Using a Simulated Radar Task, Report no. FAA-AM-78-11.

[27] BOONE, J.O., VAN BURSKIRK, L., STEEN, J., (Sept. 1980), "The Federal Aviation Administrations", Radar Training Facility and Employees Selection and Training, Report no. FAA-AM-80-15.

[28] TRACKRAY, R.I., TOUCHSTONE, R.M., "An Exploratory Investigation of Various Assessment Instruments as Correlates of Complex Visual Monitoring Performance".

[29] TRACKRAY, R.I., TOUCHSTONE, R.M. (April 1982), "Performance of Air Traffic Control Specialists (ATC's) on a Laboratory Radar Monitoring Task: An Exploratory Study of Complacency and a Comparison of ATCS and Non-ATCS Performance", Report no. FAA-AM-82-1.

[30] THACKRAY, R.I., TOUCHSTONE, R.M., (January 1988), "An Evaluation of the Effects of High Visual Taskload on the Separate Behaviors Involved in Complex Monitoring Performance", Office of Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma City, Report no. FAA-AM-88-1.

[31] THACKRAY, R.I., TOUCHSTONE, R.M., (March 1986), "Complex Monitoring Performance and the Coronary-Prone Type A Behavior Pattern", Office of Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma City, Report no. FAA-AM-86-4.

[32] MELTON, C.E., WICKS, M., (August 1967), "In-Flight Physiological Monitoring of Student Pilots", Office of Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma City, Report no. FAA-AM-67-15, 12 pp.

[33] VAUGHAN, J.A., HIGGINS, E.A., FUNKHOUSER, G.E., GALERSTON, E.M., (May 1968), "The Effects of Body Thermal State on Manual Performance", Office of Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma City, Report no. FAA-AM-68-13, 10 pp.

[34] MELTON, C.E., HOFFMANN, M., DELAFIELD, R.H., (July 1969), "The Use of a Tranquilizer (Chlordiazepoxide) in Flight Training", Office of Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-69-12, 6 pp.

[35] BILLINGS, C.E., WICK, R.L., GERKE, R.J., CHASE, R.C., (January 1972), "The Effects of Alcohol on Pilot Performance During Instrument Flight", Office of Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-72-4, 74 pp.

[36] IAMPIETRO, P.F., MELTON, C.E., HIGGINS, E.A., VAUGHAN, J.A., HOFFMANN, S.M., FUNKHOUSER, G.E., SALDIVAR, J.T., (May 1972), "High Temperature and Performance in a Flight Task Simulator", Office of Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. AFF-AM-7217, 8 pp.

[37] REVZIN, A.M., "Effects of Ethanol on Visual Unit Activity in the Thalamus", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-78-2.

[38] PEARSON, D.W., THACKRAY, R.I., (April 1969), "Consistency of Performance Change and Autonomic Response as a Function of Expressed Attitude Toward a Specific Stress Situation", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-69-7, 7 pp.

[39] YANOWITCH, R.E., STANLEY, R.M., NICHOLS, E.A., (January 1972), "Psychosocial Reconstruction Inventory - a Postdictal Instrument in Aircraft Accident Investigation", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-72-2.

[40] YANOWITCH, R.E., BERGIN, J.M., NICOLS, E.A., (March 1973), "The Aircraft as an Instrument of Self Destruction", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-73-5, 5 pp.

[41] GERATHEWOHL, S.J., (April 1978), "Psychophysiological Effects of Aging - Developing a Functional Age Index for Pilots: II. Taxonomy of Psychological Factors", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-78-16, 74 pp.

[42] JENNINGS, A.E., CHILES, W.D., WEST, G., (May 1972), "Methodology in the Measurement of Complex Human Performance: Two-Dimensional Compensatory Tracking", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-72-21, 11 pp.

[43] LEWIS, M.F., (April 1973), "Frequency of Anti-Collision Observing Responses by Solo Pilots as a Function of Traffic Density, ATC Traffic Warnings, and Competing Behavior", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-73-6, 4 pp.

[44] HASBROOK, A.H., RASMUSSEN, P.G., (June 1973), "In-Flight Performance of Civilian Pilots Using Moving-Aircraft and Moving-Horizon Attitude Indicators", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-73-9, 17 pp.

[45] LEWIS, M.F., MERTENS, H.W., (September 1978), "Pilot Performance During Simulated Approaches and Landings made with Various Computer-Generated Visual Glidepath Indicators", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-79-4, 56 pp.

[46] COLLINS, W.E., CHILES, W.D., (October 1978), "Laboratory Performance during Acute Intoxication and Hangover", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-79-7, 28 pp.

[47] COLLINS, W.E., (October 1979), "Performance Effects of Alcohol Intoxication and Hangover at Ground Level and at Simulated Altitude", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-79-26.

[48] MCKENZIE, J.M., HIGGINS, E.A., FOWLER, P.R., FUNKHOUSER, G.E., WHITE, M.A., MOSER, E., (January 1983), "Sensitivity of Some Tests for Alcohol Abuse: Findings in Nonalcoholics Recovering from Intoxication", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-83-2.

[49] COLLINS, W.E., MERTENS, H.W., (January 1988), "Age, Alcohol, and Simulated Altitude: Effects on Performance and Breathalyzer Scores", Office Aviation Medicine, Civil Aeromedical Research Institute, Oklahoma city, Report no. FAA-AM-88-2.

[50] "Formation professionnelle des conducteurs de véhicules commerciaux", (1974). Cahier de l'IRU, no. 2.

[51] "Équipements informatiques actuels et futurs dans le transport routier en Europe", (1975). Cahier de l'IRU, no. 7.

[52] "La formation du conducteur professionnel-Manuel pour les instructeurs: Transports de Marchandises", (1986). Document pratique de l'IRU.

[53] "La formation du conducteur professionel-Manuel pour les instructeurs: Transports de personnes, (1988). Document pratique de l'IRU.

[54] BERGERON, J., VAUGEOIS, P., "Performance impairements as a function of reduced perceptual or cognitives capabilities of vehicle operators", technical report, C.R.T. Université de Montréal, to appear (1990).

[55] JOLY, P., BERGERON, J., (1989), "Étude psychophysiologique de la prise de risque dans une tâche de conduite simulée", Compte rendus de la VIème conférence canadienne multidisciplinaire sur la sécurité routière. Fredericton, New-Brunswick.

[56] WILDE, G.J.S., (1982), "The theory of risk homeostasis: implications for safety and health", Risk analysis, 2, 209-225.

[57] WICKENS, C.D., (1984), "Processing ressources in attention", in R. Parasuraman, D.R. Davies (Eds.): Varieties of attention, New-York: Academic press, inc.

[58] VAUGEOIS, P., (1990), "Étude des relations entre la charge mentale de travail et la performance dans une tâche de conduite automobile simulée", Ph.D. dissertation project directed by J. Bergeron, Psychology department, Université de Montréal.

[59] TAYLOR, D.H., (1964), "Driver's galvanic skin response and the risk of accident", Ergonomics, 7, 439-451.

[60] BERGERON, J., AUDET, T., VAUGEOIS, P., "Valeurs et limites des tests d'attention pour mesurer les différences individuelles de performance dans la conduite d'un véhicule", technical C.R.T., Report, Université de Montréal, to appear (1990).

[61] BROADBENT, D.E., (1958), "Perception and Communications", London Pergamon Press.

[62] BROADBENT, D.E., (1982), "Task Combination and Selective intake of information", Acta Psychologica, 50, pp. 253-290.

[63] TRISMAN, A., (1964), "Verbal, language and meaning in selective attention", American journal of psychology, 77, pp. 206-219.

[64] TRISMAN, A., (1969), "Strategies and Models of Selective attention", Psychological Review, 76, pp 282-299.

[65] LABIALE, G., CLAIRET, J.M., AUPETIT, J., VALLET, M., (1984), "Analyse du comportement du conducteur automobile", relations entre variables; styles de conduite et consommation de carburant", CERNE, working paper, Cédex, France.

[66] CATTEL, R., EBER, H.N., (1971), Manuel du test 16 pf. Editions du Centre de Psychologie Appliquée, 41 pp.

[67] ROSENZWEIG, S., (1977), "Guide for Research on Rosenzweig Frustration Test 1934-1974", revue de Psychologie Appliquée, vol. 27, no. 1, pp. 51-61.

[68] LOPEZ, M.Y., (1963), "Le psychodiagnostic myokinétique seconde édition"- Edition du Centre de Psychologie Appliquée, 159 pp.

[69] (SEE LABIALE ET AL.)

[70] ROTTER, J.B., (1966), "Generalized expectencies for internal versus external control of reinforcement", Psychological Monographs, 80 (Whole n° 609).

[71] KAGAN, J., (1965), "Matching Familiar Figures Test", Cambridge, Massachussetts, Harvard University.

[72] HILAKIVI, I., VEILAHTI, J., ASPLUND, P., SINIVUO, J., LAITINEN, L., (1989), "A Sixteen-Factor Personality Test for Predicting Automobile Driving Accidents of Young Drivers", Acc. Anal & Prev. vol. 21, no. 5, pp. 413-418.

[73] SHAW, M., (1980), "Tachographs (use in the Trucking Industry)", On the Road, vol. 3, no. 1, pp. 19-21.

[74] COWSIN, A., (1983), "Putting a stop to speeding", Coaching Journal and Bus Review, vol. 51, no. 12.

[75] "Construction d'un simulateur de conduite", chronique du C.R.T., routes et transports, vol. XIX, (no. 2), 1989.

[76] "Le TT150- Un progrès dans la formation des conducteurs", IRU communication, CD/3753, November 1988.

[77] Campbell Scientific Canada Corporation, Newsletter vol. 4, 1987.

[78] DOBIAS, G., (1988), "Prometheus Résultats de la première étape de définition", Recherche Transports Sécurité, no. 1, pp. 52-54.

[79] FRENCH, R.L, (1989), "The Roles of Cooperative Programs in Developing Vehicular Navigation and Route Guidance Systems", pp. 7-12.

[80] KELLER, H., (1989), "Comprehensive Evaluation of Advanced Technologies in Transportation Engineering", pp. 251-256.

[81] FIALA, E., (1988), "Automobiltechnische Zeitschrift".

[82] KRAMER, U., MARX, D., POVEL, R., ZIMDAHL, W., (1987), "Automobiltechnische Zeitschrift".

[83] SEIFFERT, U., (1987), "The car and road transport, The Prometheus project".pp. 200-206.

8. LIST OF ORGANIZATIONS REACHED

CABANO TRANSPORTS (Trucking Company)

CBC : Canadian Broadcasting Corporation

FAA : Federal Aviation Administration

HUNGARO CAMION (Trucking Company)

ICAO : International Civil Aviation Administration

IMO : International Maritime Organization

IRU : International Road Transport Union

L'organisation des Pilotes de la voie maritime du fleuve St-Laurent

OTIF : Office central des transports internationaux

PROST TRANSPORT (Trucking Company)

STCUM : Société de transports de la communauté urbaine de Montréal

9. LIST OF INDIVIDUALS CONTACTED

Professor B. Aldman, Mechanical Engineering Department, Chalmers University, Sweden.

Professor J. Bergeron, Psychology Department, Université de Montréal, Montréal.

Mr. I. Costa, producer C.B.C., Toronto, Canada.

Captain R. Flores, oil cargo ship captain, president of the Marina Mercante de Venezuela, Caracas, Venezuela.

Dr. S. Finkelstein, director, medical section, ICAO, Montréal, Canada.

Mr. M. Goltran, secrétaire, IRU, Geneva, Switzerland.

Mr. O. Kervella, administrative technical officer, cargoes section, IMO, London, England.

Dr. Köfalvi, director, HUNGAROCAMION, Budapest, Hungary.

Dr. A. Lerons, first secretary, OTIF, Paris, France.

Ms. N. Marcelais, Chief, explotation Methods for Metro and Trains, S.T.C.U.M., Montréal, Canada.

Dr. K. Nedd, consultant, Applied Brain Behavior Systems, Vancouver, Canada.

M. M. Prost, director, PROST TRANSPORTS, SA, Cedex, France.

Dr. D. Suzuki, C.B.C., Toronto, Canada.

Captain J. Thompson, head of Navigation section, OMI, London, England.

Dr. R.I. Thackray, researcher, FAA, Oklahoma, U.S.A.

M. P. Vaugeois, psychologist, Ph.D. student at the Psychology Department, Université de Montréal.

10. ACKNOWLEDGEMENTS

The authors would like to thank the organizations and invididuals contacted and to acknowledge the dedication of Ms. Lucie Cournoyer who typed the manuscript.